Information theory: Coding

Yuval Filmus

December 24, 2017

Consider a scenario in which agent A wants to transmit a message of length n bits to agent B. The problem is that the message is transmitted through a channel which flips each bit with probability $p < 1/2$ independently. How should A encode her message so that B is able to decode it successfully with probability $1 - \epsilon$?

1. **Rate upper bound.** For the upper bound on the rate of communication, we assume that the message is chosen randomly, and we only require the communication to succeed with probability $1 - \epsilon$ (where the randomness is over both message and noise). We use the following notation:

   - X is the message that A wants to transmit. We assume that X is chosen uniformly at random from $\{0, 1\}^n$.
   - M is the encoded message, which we assume is of length m bits.
   - Y is the encoded message after passing through the channel. That is, $\Pr[Y_i = M_i] = 1 - p$ and $\Pr[Y_i = 1 - M_i] = p$, where $p < 1/2$ and the channel acts independently on $X_1, \ldots, X_m$.
   - Z is the decoded message. We are guaranteed that $\Pr[X = Z] \geq 1 - \epsilon$, where $\epsilon \leq 1/2$.

   We assume that M is a deterministic one-to-one function of X and that Z is a deterministic function of Y (deterministic means *not randomized*).

   (a) Let E be the indicator variable for the event "$X = Z$" (so $E = 1$ if $X = Z$ and $E = 0$ if $X \neq Z$). Show that $H(X|Z, E) \leq \epsilon n$.

   (b) Deduce that $H(M|Y) = H(X|Y) \leq H(X|Z) \leq h(\epsilon) + \epsilon n$ and so $I(M; Y) \geq (1 - \epsilon)n - h(\epsilon)$.

   (c) Show that $H(Y|M) = h(p)m$ and so $I(M; Y) \leq (1 - h(p))m$.

   (d) Let $\epsilon(n)$ be a function tending to zero, and let $m(n)$ be the minimum value of m for which such an encoding exists for the error parameter $\epsilon(n)$. Show that

   $$\limsup_{n\to\infty} \frac{n}{m(n)} \leq 1 - h(p).$$

   The quantity n/m is known as the *rate* of the encoding scheme.

2. **Rate lower bound.** Fix parameters $p < 1/2$ and $\delta > 0$.
Given n, let $m = (n+1)/(1 - h(p+\delta) - \delta)$ (we assume for simplicity that m is an integer). We will construct an encoding of n-bit messages by m-bit codewords that is able to withstand the channel which flips each of the m bits with probability p independently.

   (a) Using the law of large numbers (or directly, using Chebyshev's inequality), show that the probability that the channel flips more than $(p+\delta)m$ bits tends to zero with n.

   (b) Suppose that we choose $c_1, \ldots, c_{2^{n+1}} \in \{0,1\}^m$ uniformly at random. For each i, let B_i consist of all vectors at Hamming distance at most $(p+\delta)m$ from c_i. Show that $|B_i| \le 2^{h(p+\delta)m}$, and deduce that for every $x \in \{0,1\}^m$, the probability (over the choice of $c_1, \ldots, c_{2^{n+1}}$) that $x \in B_1 \cup \cdots \cup B_{2^{n+1}}$ is at most $2^{-\delta m}$.

   (c) Fix $1 \le i \le 2^{n+1}$. Let M_i be obtained by flipping each bit of c_i with probability p independently, and let E be the event that at most $(p+\delta)m$ bits were flipped. Let $P_i = \Pr[\exists j \neq i \text{ s.t. } M_i \in B_j \mid E]$ be the probability that $M_i \in B_j$ for some $j \neq i$ given that E happened. Show that the expected value of P_i (over the choice of $c_1, \ldots, c_{2^{n+1}}$) is at most $2^{-\delta m}$. (Hint: Express P_i as the expectation over c_i of the expectation over c_j for $j \neq i$ of an indicator variable, and use a slight modification of (b).)

   (d) Show that there exists a choice of $c_1, \ldots, c_{2^{n+1}}$ such that $\frac{P_1 + \cdots + P_{2^{n+1}}}{2^{n+1}} \le 2^{-\delta m}$.

   (e) Rearrange the c_i so that $P_1 \le P_2 \le \cdots \le P_{2^{n+1}}$. Show that $P_1, \ldots, P_{2^n} \le 2^{-(\delta m - 1)}$.

   (f) Consider the following communication system:

      - Encoding: Given a message $x \in \{0,1\}^n$, transmit $\mu = c_{x_0 + 2x_1 + \cdots + 2^{n-1}x_{n-1}}$.
      - Decoding: Given a message $y \in \{0,1\}^m$, find the minimum i such that $y \in B_i$, and decode i to a message in $z \in \{0,1\}^n$. If no such i exists, return an arbitrary message.

      Suppose that y is obtained from μ by flipping each bit with probability p independently, and let $\epsilon = \max_{x \in \{0,1\}^n} \Pr[z \neq x]$. Show that ϵ tends to zero with n (you have to consider two sources of error).

   (g) Show that for every $\gamma > 0$ we can choose $\delta > 0$ so that n/m tends to $1 - h(p) - \gamma$.

3. Explain the following sentence and its significance:
The upper bound shows that the rate is at most $1 - h(p)$ even if the communication system only needs to succeed on average, whereas the lower bound shows that every rate strictly below $1 - h(p)$ is achievable even if we require the communication system to succeed in the worst case.

Bonus question: The channel we have considered above is known as the binary symmetric channel $BSC(p)$. Another common channel is the binary erasure channel $BEC(p)$, which changes each of the message bits to a new symbol # with probability p independently. What is the optimal transmission rate of this channel?